Are Options on Index Futures Profitable for Risk Averse Investors? Empirical Evidence

George M. Constantinides  
University of Chicago and NBER

Michal Czerwonko  
Concordia University

Jens Carsten Jackwerth  
University of Konstanz

Stylianos Perrakis  
Concordia University

Abstract
American options on the S&P 500 index futures that violate the stochastic dominance bounds of Constantinides and Perrakis (2007) over 1983-2006 are identified as potentially profitable investment opportunities. Call bid prices more frequently violate their upper bound than put bid prices do, while evidence of underpriced calls and puts over this period is scant. In out-of-sample tests of stochastic dominance, inclusion of short positions in such overpriced calls and puts in the market portfolio increases the expected utility of any risk averse investor, net of transaction costs and bid-ask spreads. The results are robust and strongly supportive of economically relevant mispricing and of the stochastic dominance bounds as identifiers of mispriced options.

Current draft: November 12, 2009
(JEL G11, G13, G14)

Keywords: option mispricing; futures options; derivatives pricing; stochastic dominance; transaction costs; market efficiency

We thank Oleg Bondarenko, Wolfgang Buehler, Jim Hodder, Robert Merton, Ingmar Nolte, Myron Scholes, Sorin Sorescu, Suresh Sundaresan, Michael Wolf, participants at the Second Mont Tremblant Risk Management conference, the ESSFM Gerzensee 2008 conference, the 2008 Conference on Financial Innovation at Vanderbilt University, the 2008 International Conference on Price, Liquidity, and Credit Risks at Konstanz University, seminars at Mannheim University, Tel Aviv University, the University of Cyprus, and Zurich University, and especially Russell Davidson for insightful comments and constructive criticism. We remain responsible for errors and omissions. Constantinides acknowledges financial support from the Center for Research in Security Prices of the University of Chicago; Czerwonko and Perrakis acknowledge financial support from the Social Sciences and Humanities Research Council of Canada. E-mail addresses: gmc@ChicagoGSB.edu, mczerwon@jmsb.concordia.ca, Jens.Jackwerth@uni-konstanz.de, SPerrakis@jmsb.concordia.ca
We identify American call and put options on the S&P 500 index futures over 1983-2006 that violate the stochastic dominance bounds of Constantinides and Perrakis (2007) as potentially profitable investment opportunities. We then consider the profits that accrue from the exploitation of such mispricing by adopting the appropriate trading policy for a generic investor holding only the index and the riskless asset. In both the identification of mispriced options and the trading policy we recognize realistic trading conditions by using only observable information and by incorporating transaction costs, bid-ask spreads, and trading delays by waiting one quote before entering the position.

We show that trading policies that exploit these violations lead to out-of-sample portfolio returns that stochastically dominate (in the second order) portfolio returns that do not exploit these violations. This means that the expected utility of any risk averse generic investor increases when exploiting these violations, independent of the investor’s particular endowment and risk-averse utility function. In particular, preference for skewness is ruled out as a potential explanation for the observed mispricing.

The novel feature of our paper is that we assess the profitability of our trading policy by employing the powerful statistical tests of stochastic dominance by Davidson and Duclos (2000, 2006) which can deal with option returns even in a setting where we make minimal assumptions about investor preferences and portfolio return distributions. The comparisons are valid from the perspective of any risk averse investor, not just from the perspective of a mean-variance investor. These tests compare the profitability of the optimal trading policies of a generic S&P 500 index investor with and without the option, in a setting that recognizes the possibility of early exercise of the futures option.

We use the Chicago Mercantile Exchange (CME) database on S&P 500 futures options, 1983-2006, which is clean and spans a long period. Much of the earlier empirical work on the
mispricing of index options is based on data on the S&P 500 index options that comes from two principal sources: the Berkeley Options Database (1986-1995) that provides relatively clean transaction prices, but misses important events over the past 14 years, such as the 1998 liquidity crisis, the dot-com bubble, and its 2000 burst; and the OptionMetrics (starting 1996) database which, however, is of uneven quality and contains only end-of-day quotes.

We identify mispriced options by applying the stochastic dominance bounds for American options derived by Constantinides and Perrakis (2007). These bounds identify reservation purchase and reservation write prices such that any risk averse investor may increase her expected utility by including the option that violates these bounds in her portfolio. In the derivation of the bounds the investor is restricted to holding a combination of the index and the riskless asset. Nonetheless, we show in our tests that including the mispriced option in the investor portfolio may increase the expected utility of an investor holding a portfolio that includes the index, the riskless asset, and possibly other assets as well. Furthermore, the bounds are valid for any distribution of the underlying asset, including the empirical ones extracted from past data. The bounds are thus shown to be useful in identifying utility increasing trades beyond the restrictive set of assumptions under which they were derived.

Ample evidence supports the assumption that there exists a class of traders holding portfolios containing only the S&P 500 index and the risk free asset. Surveys report that a large number of US investors follow indexing policies in their investments. Bogle (2005) reports that, in 2004, index funds account for about one third of equity fund cash inflows since 2000 and represent about one seventh of equity fund assets. The S&P 500 index is not only the most widely quoted market index, but has also been available to investors through exchange traded

---

1 In the absence of transaction costs, Oancea and Perrakis (2009) show that the bounds encompass the well-known option pricing models under diffusion, jump-diffusion, and stochastic volatility asset dynamics, provided that they do not include additional state variables.
funds for several years. We find that any such investor would improve her utility by including in her portfolio an option identified as mispriced by the stochastic dominance bounds.

Finally, both the bounds employed in detecting mispriced options and the optimal trading policies with and without the option explicitly take into consideration bid-ask spreads, trading costs and trading delays. In our case, once a trading opportunity is detected, we execute the trade by buying at the next ask price or selling at the next bid price. Note that we are only using the bound violations as signals for trading. Thus, the assumptions underlying the bounds might even be violated as long as the signals are sufficiently accurate for profitable trading.

Our tests are non-parametric in the sense that we do not assume any particular distribution for the underlying asset returns. Therefore, our finding of mispriced options cannot be attributed to stochastic volatility and jumps in the index price. We use historical data on the underlying S&P 500 index returns in order to estimate the bounds. We use several empirical estimates of the underlying return distribution, all of them observable at the time the trading policy is implemented. For each one of these estimates, we evaluate the corresponding bounds over the period 1983-2006 and then identify the observed S&P 500 futures options prices that violate them. For each violation, we identify the optimal trading policy of a generic investor with and without the mispriced option, using the observed path of the underlying asset till option expiration and recognizing realistic trading conditions such as possible early exercise and transaction costs. We identify the profitability of the pair of policies for each observed violation and then conduct several stochastic dominance tests over the entire sample of violations.

We find a substantial number of violations of the upper bounds, but relatively few violations of the lower bounds. Since the frequency of violations of the lower bounds is too low for statistical inference, we focus on violations of the upper bounds. In all our tests, we find that
the portfolio that includes the mispriced options dominates the portfolio without the options. The dominance holds for all empirical estimates of the underlying asset distribution and is robust to the relaxation of several assumptions about the investor and the market. In particular, we allow the investor to hold a portfolio that includes other assets, beyond the index and a riskless bond, with little effect on the results of the dominance tests.

The results are thus strongly supportive of mispricing. The results also demonstrate the ability of the stochastic dominance bounds to identify mispriced options, in contrast to the failure of a naïve heuristic which we based on observed option price percentiles for buying low and selling high.²

A large body of finance literature addresses the mispricing of options. Rubinstein (1994) and Jackwerth and Rubinstein (1996) observe a steep index smile in the implied volatility of S&P 500 index options that suggests that out-of-the-money (OTM) puts are too expensive. Indeed, a common hedge-fund policy is to sell OTM puts. Coval and Shumway (2001) find that buying zero-beta, at-the-money (ATM) straddles/strangles loses money. Santa-Clara and Saretto (2009) also find that strategies of selling index options are good deals. The results of Ait-Sahalia and Lo (2000), Jackwerth (2000), and Rosenberg and Engle (2002) are suggestive of stochastic dominance, albeit in frictionless markets with a representative agent.³ By contrast, our results on stochastic dominance allow for both an incomplete market (thereby removing the requirement that a representative investor exists) and frictions.

² Although there are not enough observations to verify the quality of the lower bounds as identifiers of undervalued options, we demonstrated this ability by allowing the investor to hold in her portfolio artificial options with values equal to the lower bound. The results, reported in the on-line appendix of the paper, show the dominance of the portfolio that includes the artificial options over the portfolio without the options.

³ The assumption of a representative agent can be justified if the market is complete. Ait-Sahalia and Lo (2000, pages 25-26) call for extensions which do not rely on complete markets.
Constantinides et al. (2009) provide empirical evidence that both European puts and calls on the S&P 500 index are mispriced by showing that they violate corresponding stochastic dominance bounds on European options put forth by Constantinides and Perrakis (2002). Constantinides et al. (2009) estimate the time-series process of the index price, use this process to calculate upper and lower stochastic dominance bounds on the prices of options, and report the observed violations of the bounds by option prices. This process is subject to model misspecification and is estimated with error. Therefore, the bounds are calculated with error. The reported violations do not account for this error.

We address the above concerns in this paper as follows. The claim of stochastic dominance is not based on the observation of option prices violating the bounds. Observed violation of the bounds only triggers the trading of these options. The claim of stochastic dominance is based on the statistical test whether the portfolios that incorporate such options stochastically dominate portfolios which do not incorporate them. Even if all the assumptions that lead to the theoretical development of the bounds do not hold in practice, this does not detract from the finding of stochastic dominance based on the empirical tests but, instead, makes the claim of stochastic dominance conservative and implies that the estimated bounds may still be used as identifiers of mispriced options.

Unlike the results in Constantinides et al. (2009), our tests of stochastic dominance do allow for error in the realized returns. The statistical tests are based on the one-month realized distribution of returns of these portfolios. The tests are non-parametric and, therefore, free from assumptions regarding the return distribution. The tests are not based on the estimated time-series process of the index price and, therefore, are free from estimation error and model misspecification of the time-series process of the index price. The reported findings of stochastic
dominance are conservative because potential errors in calculating the bounds result in a trading rule that is less efficient in spotting mispriced options. Finally, we estimate the bounds at time $t$ based only on information available at time $t$. Therefore, both the calculation of the bounds and the tests of stochastic dominance are truly out-of-sample.

The paper is organized as follows. In Section I, we present the restrictions on futures option prices imposed by stochastic dominance and discuss the underlying assumptions. In Section II, we describe the data and the empirical design. In Section III, we present the empirical results and investigate their robustness. In Section IV, we discuss the implications of our results and conclude.

I. Restrictions on Futures Option Prices Imposed by Stochastic Dominance

We summarize the model and assumptions in Constantinides and Perrakis (2007) that lead to the bounds that signify violations of stochastic dominance. We stress that, even if these assumptions do not hold in practice, this does not detract from the finding of stochastic dominance based on the empirical tests reported in this paper, but instead makes the claim of stochastic dominance conservative: the tests of stochastic dominance do not depend on the assumptions made in deriving the bounds.

We allow the market to be incomplete and market agents to be heterogeneous. We investigate the restrictions on option prices imposed by one particular class of agents that we simply refer to as “traders”. We allow for other agents to participate in the market but this allowance does not invalidate the restrictions on option prices imposed by the traders.
We consider a market with several types of financial assets. First, we assume that traders invest only in two of them, a bond and a stock with the natural interpretation as a market index. Subsequently, we assume that traders can invest in a third asset as well, an American call or put option on the index futures. The bond is risk free and has total return $R$. The stock has ex-dividend stock price $S_t$ at time $t$ and pays cash dividend $\gamma S_t$, where the dividend yield $\gamma$ is deterministic. The total return on the stock, $(1 + \gamma)(S_{t+1}/S_t)$, is assumed to be i.i.d. with mean $R_s$. The call or put option on the index futures has strike $K$ and expiration date $T$. The underlying futures contract is cash-settled and has maturity $T^F$, $T^F \geq T$. We assume that the futures price $F_t$ is linked to the stock price by the approximate cost-of-carry relation

$$F_t = (1 + \gamma)^{(T^F - t)} R^{T^F - t} S_t + \epsilon_t, \quad t \leq T^F, \quad |\epsilon_t| \leq \bar{\epsilon},$$

where the basis risk $\epsilon_t$ is serially independent and independent of the stock price.

Transfers to and from the cash account (bond trades) do not incur transaction costs. Stock trades decrease the bond account by transaction costs equal to the absolute value of the dollar transaction, times the proportional transaction costs rate, $k$, $0 \leq k < 1$. Transaction costs, exchange fees, and price impact are accounted for in what we refer to as the bid and ask prices of options.

We assume that traders maximize generally heterogeneous, state-independent, increasing, and concave utility functions. We further assume that each trader’s wealth at the end of each period is weakly monotone increasing in the stock return over the period. For example, a trader who holds 100 shares of stock and a net short position in 200 call options violates the

---

4 Essentially, we model buy-and-hold investors who trade infrequently and thus incur low transaction costs. At least for large investors who earn a fair return on their margin, transaction costs are even lower in the index futures market than the stock market. In practice, however, buy-and-hold investors invest in the stock and bond markets because of the inconvenience, cost, and basis risk of the frequent rolling over of short-term futures contracts and the illiquidity of long-term futures and forward contracts.
monotonicity condition, while a trader who holds 200 shares of stock and a net short position in 200 call options satisfies the condition. Essentially, we assume that the traders have a sufficiently large investment in the stock, relative to their net short position in call options (or, net long positions in put options), such that the monotonicity condition is satisfied.5

We do not make the restrictive assumption that all market agents belong to the class of utility-maximizing traders. Thus, our results are robust and unaffected by the presence in the market of agents with beliefs, endowments, preferences, trading restrictions, and transaction costs schedules that differ from those of the utility-maximizing traders modeled in this paper.

A trader enters the market at time zero with \( x_0 \) dollars in bonds and \( y_0 \) dollars in \textit{ex dividend} shares of stock. We normalize the stock (or, index price) to \( y_0 \) dollars so that the trader holds one share (or, one unit of the index). We consider two scenarios. In the first scenario, the trader may trade the bond and stock but not the options. The trader makes sequential investment decisions at discrete trading dates \( t (t = 0, 1..., T') \), where \( T', T' \geq T^F \geq T \), is the finite terminal date. The trader’s objective is to maximize expected utility, \( E[u_r(W_T)] \), where \( V_T \) is the trader’s net worth at date \( T' \). Utility is assumed to be concave and increasing and defined for both positive and negative terminal worth, but is otherwise left unspecified. We refer to this trader as the \textit{index} (and bond) trader, IT, and denote her maximized expected utility by \( V^{IT}_{0}(x_0, y_0) \).

In the second scenario, as in the first scenario, the trader enters the market at time zero with \( x_0 \) dollars in bonds and \( y_0 \) dollars in \textit{ex dividend} shares of stock, but immediately writes one American futures call option with maturity \( T, T \leq T^F \), where \( C \) are the net cash proceeds.

---

5 The monotonicity condition must hold only within each trading period; it need not hold for the return of the index plus option portfolio over the entire holding period to option expiration.
from writing the call.\textsuperscript{6,7} We assume that the trader may not trade the call option thereafter.\textsuperscript{8} At each trading date \( t \) (\( t = 0, 1 \ldots , T \)) the trader is informed whether or not she has been \textit{assigned} (that is, assigned to act as the counterparty of the holder of a call who exercises the call at that time). If the trader has been assigned, the call position is closed out, the trader pays \( F_t - K \) in cash, and the value of the cash account decreases from \( x_t \) to \( x_t - (F_t - K) \). The trader makes sequential investment decisions with the objective to maximize her expected utility, \( W = \sum_{t=0}^{T} u_t(W_t) \).

We refer to this trader as the \textit{option} (plus index and bond) trader, OT, and denote her maximized expected utility by \( V_0^{\text{OT}}(x_0 + C, y_0) \).

For a given pair \((x_0, y_0)\), we define the reservation write price of a call as the value of \( C \) such that \( V_0^{\text{OT}}(x_0 + C, y_0) = V_0^{\text{OT}}(x_0, y_0) \). The interpretation of \( C \) is the write price of the call at which the trader with initial endowment \((x_0, y_0)\) is indifferent between writing the call or not.

Constantinides and Perrakis (2007) state a tight upper bound on the reservation write price of an American futures call option that is independent of the trader’s utility function and initial endowment and independent of the early exercise policy on the calls:

\textsuperscript{6} We normalize the size of a futures contract to be on one unit of the index; and we normalize the size of the futures option to be on one futures contract.
\textsuperscript{7} The reservation write price of a call is derived from the perspective of a trader who is marginal in the index, the bond, and only one type of call or put option at a time. Therefore, these bounds allow for the possibility that the options market is segmented.
\textsuperscript{8} The reservation write price of a call is derived under this constrained policy. Define \( \overline{C} \) as an upper bound on the reservation write price of a call. Under this constrained policy, the investor increases her expected utility by writing a call at price \( \overline{C} \) and refraining from trading the call thereafter. If the constraint on trading the call is relaxed, the policy which the investor follows under the constraint policy remains feasible and increases her expected utility by writing a call at price \( \overline{C} \). Therefore, \( \overline{C} \) remains an upper bound on the reservation write price of a call. Whereas the upper bound may be tightened when the constraint on trading the call is relaxed, there is no known tighter bound that is preference free. For further discussion on these bounds, see Constantinides and Perrakis (2007).
The function $N(S,t)$ is defined as follows:

$$N(S,t) = \left( R_s^{-1} \right) E \left[ \max \left\{ \left( 1 + \gamma \right)^{\left( T - t - 1 \right)} R^T - t - 1 S_{t+1} + \bar{\gamma} - K, N(S_{t+1}, t+1) \right\} \bigg| S_t = S \right], \quad t \leq T - 1$$

$$= 0, \quad t = T.$$  \hspace{1cm} (2)

The economic interpretation of the call upper bound is as follows. If we observe a call bid price above the reservation write price, $\bar{C}$, then any trader (as defined in this paper) can increase her expected utility by writing the call.

Transaction costs on the index have only a small effect on the upper bound. Specifically, without transaction costs on the index, the upper bound is $\max \left[ N(S,t), F_i - K \right]$; with transaction costs on the index, the upper bound merely increases by the multiplicative factor $\left( 1 + k \right) / \left( 1 - k \right)$. The reason is that this particular bound is based on a comparison of the utility of an index trader to the utility of an option trader. Both traders follow the trading policy which is optimal for the index trader but is generally suboptimal for the option trader. This policy incurs very low transaction costs because the trader trades infrequently, as shown in Constantinides (1986).

If we further assume that the trader can buy a call at price $\bar{C}(F_i, S_i, t)$ or less and trade the futures and do so costlessly, we obtain the following put upper bound:

\begin{equation}
\bar{C}(F_i, S_i, t) = \frac{1 + k}{1 - k} \max \left[ N(S_i, t), F_i - K \right], \quad t \leq T. \hspace{1cm} (1)
\end{equation}

\begin{align*}
\bar{C}(F_i, S_i, t) & = \left( R_s^{-1} \right) E \left[ \max \left\{ \left( 1 + \gamma \right)^{\left( T - t - 1 \right)} R^T - t - 1 S_{t+1} + \bar{\gamma} - K, N(S_{t+1}, t+1) \right\} \bigg| S_t = S \right], \quad t \leq T - 1 \\
& = 0, \quad t = T.
\end{align*}

We prove equation (3) by noting that an investor achieves an arbitrage profit by buying a call at $\bar{C}(F_i, S_i, t)$, writing a put at $P, P > \bar{P}(F_i, S_i, t)$, selling one futures, and lending $K - R^{(T - t)} F_i$. In the proof, we ignore the daily...
\[
\overline{P}(F_t, S_t, t) = \overline{C}(F_t, S_t, t) - R^{(T-t)}F_t + K, \quad t \leq T.
\]  

with similar interpretation.\(^{10}\)

II. Data Description and Methodology

In this section, we describe our data on index, futures, and option prices. We explain how we calibrate a tree of the daily index return and use it to calculate the option bounds. We describe the construction of the portfolio of the index trader (IT) and of the option trader (OT). Finally, we explain our empirical methodology of comparing the performance of the IT and OT portfolios in terms of their means and in terms of the criterion of second order stochastic dominance.

II.A Data description and estimation

We obtain the time-stamped quotes of the 30-calendar-day S&P 500 futures options and the underlying nearest to maturity futures for the period February 1983-July 2006 from the Chicago Mercantile Exchange (CME) tapes. This results in 247 sampling dates. We obtain the interest rate as the three-month T-bill rate from the Federal Reserve Statistical Release H.15. The data sources are described in further detail in Appendix A.

marking-to-market on the futures until the exercise of the put or the options' maturity, whichever comes first. This matters little because the investor has a large investment in the bond which suffices to cover margin calls.\(^{10}\) Constantinides and Perrakis (2007) derive lower bounds on the reservation purchase price of American call and put options on futures. We do not state these lower bounds here because the observed frequency of their violation is too low for statistical inference. The methodology may be extended to individual stock options as well. The development of this theory and its empirical implementation is a promising direction for future research.
For the daily index return distribution, we use the historical sample of logarithmic returns from January 1928 to January 1983. However, when looking forward for each of our 247 option sampling dates, we adjust the first four moments of the index return distribution in various ways which we now describe in detail. We set the mean logarithmic index return at 4% plus the observed 3-month T-bill rate instead of estimating the mean index return from the data in order to mitigate statistical problems in estimating the mean.\footnote{We relax this assumption at the end of Section III. All our results are unchanged because prices of short-dated options are insensitive to the interest rate.} We implement this by adding a constant to the observed logarithmic index returns so that their sample mean equals the above target.

We estimate both the unconditional and conditional volatility of the index returns. We estimate the unconditional volatility as the sample standard deviation over the period January 1928 to January 1983.\footnote{We also estimate the unconditional volatility over the 24 years prior to January 1983. The results remain essentially unchanged and are not reported.} We estimate the conditional volatility in three different ways: (1) the sample standard deviation over the preceding 90 trading days;\footnote{We also estimate the conditional volatility over the preceding 360 days. The results remain essentially unchanged and are not reported.} (2) the ATM implied volatility (IV) on the preceding day, adjusted by the mean prediction error for all dates preceding the given date (typically some 3%);\footnote{We start with the 22\textsuperscript{nd} month. We use the holdout sample of the first 21 months to estimate the mean adjustment error and adjust the IV of the 22\textsuperscript{nd} month. We use the holdout sample of the first 22 months to estimate the mean adjustment error and adjust the IV of the 23\textsuperscript{rd} month; and so on.} and (3) the Nelson (1991) EGARCH (1, 1) model volatility using EGARCH coefficients estimated for S&P 500 daily returns over January 1928 to January 1983 applied to residuals observed over the 90 days preceding each sample date to form projections of the volatility realized till the option expiry date.\footnote{We form the volatility projections by iterating from day \(t + 1\) till the option maturity \(T\), as explained in Baillie and Bollerslev (1992). We use as inputs the past 90-day residuals and the model coefficients estimated in the pre-sample} We estimate the 3\textsuperscript{rd} and 4\textsuperscript{th} moments of the index return as their sample counterparts over the preceding 90 days.
In Table I, we report statistics of the prediction error of the above volatility estimates. The best overall predictor is the adjusted ATM IV and the second best predictor is the 90-day historical volatility.

[Table I about here]

II.B  Calibration of the index returns tree and calculation of the option bounds

We model the path of the daily index return till the option expiration on a $T$-step tree, where $T$ is the number of trading days in that particular month. The tree is recombining with $m$ branches emanating from each node. Every month we calibrate the tree by choosing the number of branches, spacing, and transition probabilities at each node to match the first four moments of the daily index return distribution, as described in Appendix B. We numerically calculate the bounds in equations (1)-(3) by iterating backwards on the calibrated tree.

II.C  Portfolio construction and trading

For each monthly stock return path, we employ the following trading policies. For the index trader (who manages a portfolio of the index and the risk free asset in the presence of transaction costs), we employ the optimal trading policy, as derived in Constantinides (1986) and extended in Perrakis and Czerwonko (2006) to allow for dividend yield on the stock. Essentially, this policy consists of trading only to confine the ratio of the index value to the bond value, $y_t / x_t$, within a no-transactions region, defined by lower and upper boundaries. We derive these boundaries for the following parameter values: one-way transaction cost rate on the index of 0.5%; annual return volatility of the index of 0.1856, the sample volatility over 1928-1983;
interest rate equal to the observed 3-month T-bill date; risk premium 4%; and constant relative risk aversion coefficient of 2.\textsuperscript{16} For this set of parameters, the lower and upper boundaries are $y_0/x_0 = 1.2026$ and $1.5259$, respectively. At the beginning of each month and before the trader trades in options, we set $x_0 = 73,300$ and $y_0 = 100,000$, which corresponds to the midpoint of the no-transactions region, $y_0/x_0 = 1.3642$.\textsuperscript{17} We normalize the index price to $y_0$ dollars so that the trader holds one unit of the index.

For the option trader (who manages a portfolio of the option, index, and the risk free asset in the presence of transaction costs), we set $x_0$ and $y_0$ to the same values as for the index trader. The option trader writes or buys one call or one put on the index futures.\textsuperscript{18} However, this portfolio composition changes, depending on the assumed position in futures options, as explained in Appendix C. We employ the trading policy which is optimal for the index trader but is generally suboptimal for the option trader. Recall that the goal is to demonstrate that there exist profitable investment opportunities for the option trader. Given this goal, it suffices to show that there exist profitable investment opportunities for the option trader even though the option trader follows a generally suboptimal policy in trading the index.

We focus on the case where the basis risk bound, $\varepsilon$, is 0.5% of the index price. Over the years 1990-2002, 95% of all observations have basis risk less than 0.5% of the index price. For

\textsuperscript{16} We clarify that the upper and lower stochastic dominance bounds on option prices apply to any risk averse trader, independent of her particular form of utility function. Only the trading policy is affected by the choice of risk-aversion, but only slightly, because the no transactions region is wide and trading is infrequent. In our empirical work, we present results for a trader with constant relative risk aversion coefficient 2 and 10.

We repeat our tests by replacing the optimal trading policy with a buy-and-hold policy. The results are virtually identical and are not reported here.

\textsuperscript{17} The results are unchanged when the starting portfolio is at either boundary of the no transactions region.

\textsuperscript{18} We normalize the size of a futures contract to be on one unit of the index; and we normalize the size of the futures option to be on one futures contract.
reference purposes, we also consider the case $\varepsilon = 0$. As expected, when we suppress the basis risk, the bounds are tighter and there appear to be more violations.

II.D Empirical methodology

For each of our methods of estimating the bounds, we obtain 247 monthly portfolio returns for the index trader and the option trader, respectively. Our goal is to test whether the portfolio profitability of the index and option traders are statistically different in the months in which we observe violations of the bounds.

We apply the criterion of second order stochastic dominance (SSD), which states that the dominating portfolio is preferred by any risk-averse trader, independent of distributional assumptions, such as normality, and preference assumptions, such as quadratic utility. Formally, the OT portfolio stochastically dominates the IT portfolio if, for every $z$ in the joint support of their respective distributions, the following holds:

$$D^2_{\text{IT}}(z) - D^2_{\text{OT}}(z) \geq 0,$$

with strict inequality for at least one value of $z$, where

$$D^2_j(z) = \int_{z}^{z} (z-x)dF_j(x), \quad J = \text{OT, IT}, \quad F_j(x) \text{ is the cumulative distribution function of the portfolio return, and } z \text{ is the lower bound of the common support.}$$
First, we test the null hypothesis \( H_0 : OT \succeq IT \) against the alternative that either \( OT \succ IT \) or that neither one of the two distributions dominates the other. Hence, rejection of the null hypothesis fails to rank the two distributions. We also test the converse null hypothesis \( H_0 : OT \succ IT \) against the alternative that either \( IT \succ OT \) or that neither one of the two distributions dominates the other. For these hypotheses, we report the results of the test proposed by Davidson and Duclos (2000) (DD (2000)), described in Appendix D. The test requires that returns be serially uncorrelated, an assumption that holds well in all our return series: the first-order serial correlation ranges from -0.0267 to 0.0964 and is statistically insignificant.

Second, we test the null hypothesis \( H_0 : OT \not\succeq IT \) which states the option trader’s portfolio return does not stochastically dominate the index trader’s portfolio return, against the alternative hypothesis \( H_A : OT \succeq IT \), which states the option trader’s portfolio return stochastically dominates the index trader’s portfolio return. Rejection of this hypothesis means that the option trader’s portfolio return stochastically dominates the index trader’s portfolio return. Likewise, we test the converse null hypothesis \( H_0 : IT \not\succeq OT \) against the alternative hypothesis \( H_A : IT \succeq OT \). For these hypotheses, we report the results of the test proposed by Davidson and Duclos (2006) (DD (2006)), using the algorithm developed by Davidson (2007). The test is described in Appendix D. Again, the test requires that returns be serially uncorrelated, an assumption that holds well in all our return series. Nolte (2008) investigates the power of the DD (2006) when there are GARCH effects and finds that the test performs well.

The power of the DD (2006) test is low, unless one trims the tails of the paired outcomes. Therefore, we trim 10% of the paired outcomes in the left tail of our sample distributions which
affects both IT and OT similarly and is therefore innocuous. Restricting the range further by trimming outcomes on the right tail of the distribution presents a problem. Without any trimming on the right tail, the test has low power. Trimming on the right, on the other hand, may bias our test towards rejection of the null, because IT tends to produce superior results to OT when the return of the underlying asset is high. For this reason, we present in our tables results with 0%, 5%, and 10% trimming on the right.\textsuperscript{19} To facilitate interpretation, we perform all our statistical tests on annualized arithmetic returns on the wealth of OT and IT investors.\textsuperscript{20}

We choose the tests DD (2000) and DD (2006) because, unlike several alternatives, they apply to correlated samples and are more powerful than other well-known tests.\textsuperscript{21}

\section*{III Empirical Results}

In Section III.A, we describe the pattern of observed violations for the bounds with respect to the degree of moneyness. In Section III.B, we present the main empirical results. We compare the portfolio return of an option trader who writes overpriced calls or puts at their bid price with the portfolio return of an index trader who does not trade in the options over the period 1983-2006. We find that the return of an option writer stochastically dominates the index trader’s return, net of transaction costs and the bid-ask spread. Whereas we find a substantial number of violations of the upper bounds, we find relatively few violations of the lower bounds. In Section III.C, we establish that the empirical results are robust. In Section III.D, we demonstrate that trading

\textsuperscript{19} In the on-line appendix of the paper, we test the effects of such trimming on simulated data that mirror our sample. Our simulations show that the test is, if anything, conservative in rejecting the false null.
\textsuperscript{20} We annualize returns since times to maturity vary from 28 to 31 days in our sample. Since transaction costs are present in our economy, we derive returns for the liquidation of the risky asset under the assumed one-way transaction costs rate of 0.5%.
\textsuperscript{21} See Tse and Zhang (2003).
policies triggered by violations of the stochastic dominance bounds consistently outperform the naïve filter rule of buying low and selling high.

III.A The pattern of violations

In Figure 1, we plot the four bounds for one-month options for May 22, 1996, expressed in terms of the implied volatility, as a function of the moneyness, $K/F$. We set $\sigma = 20\%$ and $\bar{\varepsilon} = 0$. The figure also displays the 95% confidence interval, derived by bootstrapping the 90-day distribution. The call upper bound is tighter than the put upper bound and both bounds are downward sloping. The put lower bound is tighter than the call lower bound. The put lower bound is downward sloping but the call lower bound is not.

[Figures 1 and 2 about here]

In Figure 2, we display the time pattern of actual violations of the call upper bound. The crosses display the violations of the call upper bound for the period February 1983-July 2006. For the adjusted IV distributions, the first 21 dates are not in the sample because they are needed to obtain the adjustment. The solid lines are the natural logarithm of the S&P 500 index, the VIX index, and the T-Bill rate. For all different ways of estimating volatility, we observe violations after significant down moves in the index, when we expect the implied volatility to be high.

[Table II about here]

In Table II, the violations are shown as a proportion of the quotes in each moneyness range. It is clear that, for all methods of estimating the bounds, there is a large proportion of violations. For the moneyness range 1.03-1.08, a large proportion of the available quotes violate the corresponding bound for all estimation methods.
Table III shows the violations in each moneyness range, as a proportion of the total number of quotes across the whole range of moneyness. The largest number of violations, as a proportion of the total number of quotes, is found in the 1.01-1.03 moneyness range and not in the 1.03-1.08 range because there are relatively few quotes in the latter range. For all estimation methods, a majority of the identified violations are in the liquid range, 0.99-1.03. Our data also shows that the average size of the mispricing is between 5% and 56% of the upper bound for most methods of estimating the bounds as we move from the 0.96-0.99 liquidity range to the 1.03-1.08 liquidity range. In the stochastic dominance tests, the power of the tests depends, by construction, on the proportion of months with observed violations.

We further investigate how the incidence of calls violating the upper bound relates to characteristics of the options (moneyness, ATM implied volatility, volume of trade, and put-call ratio), the index (return, momentum, presence of jumps between trading dates, dividend yield, skew, and volume of trade), the term structure, and the default spread. For each of the four volatility prediction methods, we sort the sample of calls into terciles and report the average fraction of violations in the top and bottom terciles. The results are reported in Table IV. The incidence of violations is higher when the Moody Baa-Aaa default spread is high, when the futures open interest is high, and also when momentum is low. For all other characteristics, we do not find a consistent pattern of violations because the incidence of violations depends on the prediction method of the volatility as an input to the derivation of the bounds.

III.B Empirical evidence on stochastic dominance

22 The exchange regulations specify that the minimum number of available contracts must be at least 20 for each quote.
We apply our statistical tests to all months in the sample, even though there are months in which
the OT trader does not trade in options and the returns in these months are identical for the OT
and IT portfolios. In Table V, Panels A and C, we present the cases of call and put bid prices
violating their upper bound, when we set the basis risk bound at 0.5% of the index price. (In
Panel B, we present robustness tests and discuss them in Section III.C.) We find a higher
frequency of violations of the upper call bound than of the upper put bound because the upper
call bound is tighter than the upper put bound, as we observed in Figure 1.

[Table V about here]

In our first test of stochastic dominance, we consider the hypothesis $H_0: OT \succeq_2 IT$, which states that the option trader’s return dominates the index trader’s return. We apply the DD (2000) test and obtain p-values that exceed 10% for both the upper call bound and the put upper bound. The results are not reported in the table.

In our second test, we consider the hypothesis $H_0: IT \succeq_2 OT$, which states that the index trader’s return dominates the option trader’s return. Again, we apply the DD (2000) test. In Table V, Panel A, the p-values are lower than 1% and the hypothesis is rejected when the option trader writes overpriced calls. In Panel C, the first two rows, the p-values exceed 10% but that is largely because there are very few months in which we observe violations of the put upper bound.

In our third test, we consider the hypothesis $H_0: IT \not\succeq_2 OT$ which states that either the option trader’s return dominates the index trader’s return or that neither return dominates the other. We apply the DD (2006) test and obtain p-values of one for both the upper call bound and the put upper bound. The results are not reported in the table.
Finally, we consider the hypothesis $H_0: OT \not\succ \succ IT$, which states that either the index trader’s return dominates the option trader’s return or that neither return dominates the other. As we explained earlier, the power of the DD (2006) test is low, unless we trim the tails of the paired outcomes. Therefore, we trim 10% of the paired outcomes in the left tail of our sample distributions. Without trimming on the right tail, the test has low power and the p-values are high in the two panels of Table V. Without trimming, we reject at the 10% level for one case in Panel A. In Panel A, with either 5% or 10% trimming on the right tail, the null is rejected. In Panel C, the results are inconclusive largely because there are few months in which we observe violations of the put upper bound.

Overall, the results in Table V imply that the relatively large number of violations of the call upper bound by call bid prices leads to a trading policy where the option trader’s return stochastically dominates the index trader’s return.\(^\text{23}\) The results point in the same direction for violations of the put upper bound by put bid prices, but the statistical significance is weak because there are very few months in which we observe violations of the put upper bound.

III.C Robustness tests

We now demonstrate in a number of ways that the results of Table V are robust. Table VI differs from Table V only in that the basis risk is set at zero, $\bar{\varepsilon} = 0$, instead of bounding the basis risk by $\bar{\varepsilon} = 0.5$. There are now more options across the board violating the bounds because all the bounds become tighter: the upper bounds are lowered and the lower bounds are raised. We present the cases of call and put bid prices violating their upper bound. We do not present results

\(^{23}\) We gauge the economic significance of violations of the call upper bound to be within 10 to 15% of the call bidding prices. This is the uniform proportional decrease in those prices before the stochastic dominance test results begin to deteriorate, as shown in the on-line appendix of the paper.
for the cases when the call and put ask prices violate their lower bound because we still do not have a sufficient number of such violations to be able to make statistical inference.

[Table VI about here]

Since the upper call and put bounds are lower, the options trader is less selective than before in writing options that violate their upper bounds. As in Table V, the DD (2000) test does not reject the hypothesis $H_0 : OT \succcurlyeq IT$ and rejects the hypothesis $H_0 : IT \succcurlyeq OT$. With 10% trimming, the DD (2006) test rejects the hypothesis $H_0 : OT \not\succcurlyeq IT$. We conclude that the results in Table VI are consistent with those in Table V.

As another robustness check, we set the relative risk aversion coefficient at 10 instead of 2. Since the upper and lower stochastic dominance bounds on option prices are independent of the trader’s utility, we observe the same number of violations as we do in Table V. The change in the risk aversion coefficient does change the boundaries of the no-transactions region and, therefore, the trading policy of the index trader and the option trader. The results of the stochastic dominance tests are virtually identical to those in Table V and are not reported. The results confirm our earlier conjecture that these test results are unaffected by the choice of investor risk aversion parameter. As further confirmation, we repeated the stochastic dominance tests when the trader follows a buy-and-hold policy on the index. The results are virtually identical to those in Table V and are not reported here.

As another robustness check, we set the expected premium on the index at 6% instead of 4%. Since the upper call and put bounds are higher, the options trader is more selective than before in writing options that violate their upper bounds. We also consider two other variations: setting the premium at 2%; and drawing the premium each month from a uniform distribution over 2%-6%. In all cases, the stochastic dominance results in writing calls are as strong as in
Table V.24 We conclude that the results in Table V are robust to the assumption that the expected premium on the index is 4%.

Next, we exclude from the sample the seven months from October 1987 to April 1988 in order to abstract from the crash since unusually high post-crash implied volatility might have presented unusually profitable trading opportunities. Since the stochastic dominance results are essentially the same as in Table V, we do not report these results.

Finally, we test the sensitivity of the results to the assumption that the index trader does not hold any assets other the bond and the index. We vary the composition of the IT portfolio by adding several risky assets and verify whether the modified index trader improves her utility by adding the previously identified mispriced calls. The bounds that we rely upon are no longer strictly valid because they are derived under the assumption that the index trader does not have any additional assets beyond the index and the risk free asset. Nevertheless, we rely on these bounds to find overpriced calls, thereby making it harder to find a profitable OT strategy. In Table V, Panel B, we address the case where the IT portfolio includes 20% of the “high minus low” (HML) and 20% of the “small minus big” (SMB) factors of Fama and French (1993). As before, we find that the portfolio return of the option trader who writes overpriced calls stochastically dominates the return of the modified index trader. Additional results are reported in the on-line appendix of the paper.

---

24 The results for risk premia different than 4% are presented in the on-line appendix of the paper.
III.D Comparison with naïve trading policies

In the introduction, we cited evidence that index calls and puts are generally overpriced. This motivates comparison of the portfolio return of an index trader with the portfolio return of a naïve option trader who indiscriminately writes all available calls (or, puts) on index futures every month. The results are presented in the first two lines of Table VII, Panel A. Both the DD (2000) and DD (2006) tests are weakly supportive of the hypothesis that the call option trader’s returns stochastically dominate the index trader’s returns. This conclusion is not strengthened in Panels B and C where the trader shorts the top 10% and 2.5% of the calls, respectively, because of the reduced sample size. Overall, we conclude that even indiscriminate writing of calls improves the portfolio returns in terms of the stochastic dominance criterion, although a comparison with the equivalent results of Table V shows that the evidence for the naïve policy is less powerful in all cases than the shorting of call options that violate the bounds.

[Table VII about here]

Next, we compare the portfolio return of an index trader with the portfolio return of a naïve option trader who indiscriminately writes all available puts on index futures every month. The results are presented in the second line of Table VII, Panel A. The DD (2006) test, with any amount of trimming of the right-hand tail, does not reject the hypothesis that the OT portfolio does not dominate the IT portfolio. This suggests that indiscriminate writing of puts does not improve the portfolio return in terms of the stochastic dominance criterion.

The results presented in the last two lines of Table VII, Panel A, B, and C, confirm the obvious: there is no evidence that indiscriminate buying of calls or puts leads to portfolio returns that stochastically dominate the portfolio return of the index trader.
Overall, the naïve trading rules work well in a quarter of the cases: they work well in writing calls but work poorly in writing puts, buying calls, or buying puts. Thus, the limited success of the naïve trading rules appears to be fortuitous. By contrast, the trading rules based on stochastic dominance bounds work well in practically all cases: they identify overpriced calls and puts and find very few underpriced calls and puts.

IV Concluding Remarks

We introduce a new approach for empirical research in option pricing and apply it to S&P 500 index futures options. We search for mispriced American call and put options on the S&P 500 index futures by employing stochastic dominance upper and lower bounds on the prices of options. We find a substantial number of violations of the upper bounds, but relatively few violations of the lower bounds. Since the frequency of violations of the lower bounds is too low for statistical inference, we focus on violations of the upper bounds. We observe that the highest proportion of violations occur in the region of OTM calls, where the bounds are tight. We also find, however, that the largest number of violations are in the close-to-the-money region and, hence, liable to correspond to more liquid options.

We compare the portfolio return of an option trader who writes overpriced calls or puts at their bid price with the portfolio return of an index trader who does not trade in the options over the period 1983-2006. In out-of-sample tests, our main result is that the return of a call or put writer stochastically dominates the index trader’s return, net of transaction costs and the bid-ask spread. The dominance holds under a variety of methods for estimating the underlying return distribution. It also holds when the trader is allowed to vary her portfolio position by adding other risky assets beyond the index to her portfolio.
Our results are consistent with equilibrium in a segmented market along the following lines. Mutual funds exert price pressure on OTM index puts because they buy them as insurance; and over-optimistic speculators exert price pressure on OTM calls because they buy them as a leveraged bet. Furthermore, Garleanu, Pedersen, and Poteshman (2009) argue that dealers inflate the prices of options. As we show in the paper, this presents opportunities for individual investors to write overpriced calls and puts and enhance their portfolio returns, net of transaction costs, in terms of the criterion of stochastic dominance. This can be an equilibrium if the number of such traders and the scale of their trades are sufficiently small so that they do not eliminate the overpricing. Large investors such as hedge funds who can potentially eliminate the overpricing do not write these overpriced options in a large scale because, as Santa-Clara and Saretto (2009) point out, they face obstacles including margin calls and the lack of market depth.
Appendix A: Data

S&P 500 futures have maturities only in months in the March quarterly cycle. Options on the S&P 500 futures have maturities either in a month in the March quarterly cycle (“quarterly options”) or in a month not in the March quarterly cycle (“serial options”). We consider one-month quarterly options written on one-month futures and one-month serial options written on futures with the shortest maturity. We obtain the time-stamped quotes of the one-month S&P 500 futures options and the underlying one-month futures for the period February 1983-July 2006 from the CME tapes.

From futures prices, we calculate the implied S&P 500 index prices by applying the cost-of-carry relation

\[ F_t = (1 + \gamma)(r^{t+1} - r^t) R^{t+1} S_t + \varepsilon_t , \]

assuming away basis risk, \( \varepsilon_t \equiv 0.2^5 \). We obtain the daily dividend record of the S&P 500 index over the period 1928-2006 from the S&P 500 Information Bulletin and convert it to a constant dividend yield for each 30-day period. Before April 1982, dividends are estimated from monthly dividend yields. We obtain the interest rate as the three-month T-bill rate from the Federal Reserve Statistical Release H.15. We estimate the variance of the basis risk, \( \text{var}(\varepsilon_t) \), from the observed futures prices and the intraday time-stamped S&P 500 record obtained from the CME.

We rescale the index price \( S_t \) by the multiplicative factor \( 100,000/S_0 \) so that the index price at the beginning of each 30-day period is 100,000. Accordingly, we rescale the futures price, index futures option price, and strike by the same multiplicative factor.

\[ ^{25} \text{Recall that our goal is to compare the investment policies of the index trader and the option trader. Since both policies stipulate approximately the same stock component, the effects of this component cancel each other out. Also, it is common in empirical work to derive the index value from the index futures; see, for example, Jackwerth and Rubinstein (1996).} \]
We consider options maturing in 30 calendar days, which results in 247 sampling dates. Since the first maturity of serial options was in August 1987, the first 19 periods occur with quarterly periodicity. Overall, we record 36,921 raw call quotes and 42,881 raw put quotes. After eliminating obvious data errors, we apply the following filters: minimum 15 cents for a bid quote and 25 cents for an ask quote; $K/F$ ratio within 0.96-1.08 for calls and within 0.92-1.04 for puts; and matching the underlying futures quote within 15 seconds. Part of the data is lost due to the CME rule of flagging quotes, i.e. bids (asks) are flagged only if a bid (ask) is higher (lower) than the preceding bid (ask); in addition, no transaction data is flagged. We recover a large part of the data by analyzing the sequence between consecutive bid-ask flags; however, this recovery is not possible in all cases. As a result of the applied filters, we obtain 29,822 quotes for calls and 30,281 quotes for puts in our final sample. These quantities translate into roughly 60 data points for all strikes for either bid or ask prices for an average day.

Appendix B: Calibration of the index return tree

For every month, we model the path of the daily index return till the option expiration on a $T$-step recombining tree, where $T$ is the number of trading days in that particular month. The paths of the daily index return emanate with $m$ branches from each node. The objective is to match as closely as possible the first four moments of the daily return distribution. As explained in Section II.A, we fix the mean and use the estimated volatility from one of our four methods. We use as the third and fourth moment the observed sample moments over the 90 preceding calendar days.

---

26 The 30-day rule eliminates the observation of the October 1987 crash from our sample. Therefore, we use one 40-day period for October 1987 in order to observe the crash. Our results remain unchanged.
27 For example, if the 3rd Friday of July is on July 27, we record the price of the July option on June 27, which is 30 calendar days earlier. (If June 27 is a holiday, we record the price on June 26.) If there are 21 trading days between June 27 and July 27, we model the path of the daily index return till the option expiration on a 21-step tree.
In the first step of our algorithm, we pick an odd value for the number of branches \( m \) and group the sample of daily returns in a histogram with \( m \) bins of equal length (on the log scale) such that the extreme bins are centered on the extreme observed returns. The center of each bin then becomes a state in the lattice, with the ordered states and the corresponding probabilities denoted respectively as \( x_i \) and \( p_i, \ i = 1, \ldots, m \). Note that this equidistant log-scale and an odd value for the number of branches \( m \) are necessary for the lattice to recombine.\(^{28}\)

In a second step, we match our moments by fixing the number of branches \( m \) and matching the first three moments by changing the spacing (via parameters \( a \) and \( b \)) and the probabilities via parameter \( c \). The forth moment is then matched by changing the number of branches, \( m \).

We derive the required parameters \( a, b, \) and \( c \) by solving the following set of three non-linear equations that are simply three moment conditions for the constants \( a, b, \) and \( c \):

\[
\sum_{i=1}^{m} p_i^* \exp(ax_i + b) - \exp(\hat{\mu}) = 0
\]

\[
\sum_{i=1}^{m} p_i^* \left[ \exp(ax_i + b) \right]^2 - \exp(\hat{\mu})^2 - \hat{\sigma}^2 = 0
\]

\[
\sum_{i=1}^{m} p_i^* \left[ \exp(ax_i + b) - \exp(\hat{\mu}) \right]^3 - \hat{\mu}_3 \hat{\sigma}^3 = 0
\]

where \( \exp(\hat{\mu}) \) and \( \hat{\sigma}^2 \) are the first and second target moments, respectively; \( \hat{\mu}_3 \) is the sample skewness; and \( p_i^* = \frac{p_i + c_{1 \left( \bar{x} \geq n^* \right)} \frac{1}{\bar{p}_{i \left( x \neq 0 \right)}}}{\sum_{i=1}^{m} \left( p_i + c_{1 \left( \bar{x} \geq n^* \right)} \frac{1}{\bar{p}_{i \left( x \neq 0 \right)}} \right)} \), where \( 1(\cdot) \) is the indicator function, \( n^* \) is the index to this \( x_i \) which brackets from above the target expected log-return \( \hat{\mu} \). The first indicator function ensures that the constant \( c \) is added only to the probabilities in the right tail of the

\(^{28}\) We did not build our lattice by discretizing a kernel-smoothed distribution because this method requires a substantially larger lattice. We did not adopt the Edgeworth/Gram-Charlier binomial lattice methodology, as in Rubinstein (1998), because it sometimes results in negative probabilities.
distribution; the second one ensures that the constant $c$ is added only to the positive probabilities. Note that the affine transformation of the log-states $x_i$ preserves the equal distance between the adjacent states. The constant $a$ ensures the desired scale of the log-states $x_i$, the constant $b$ ensures the desired location of these states, while the constant $c$ increases or decreases the probabilities in the right tail relative to the left one to match the desired skewness.\(^{29}\)

To match the fourth sample moment $\hat{\mu}_4$, we search over $m$, the number of nodes in the lattice. With each new $m$ the initial distribution derived from a histogram changes, providing some variability in the fourth moment after the adjustments resulting from solving (B.1). After a search over a range of $m$’s, we pick this distribution which has the lowest absolute difference between its kurtosis and the sample kurtosis $\hat{\mu}_4$. This search procedure results in very small errors in matching $\hat{\mu}_4$ for the data that we use while we obtain the exact match in the first three moments. For the four volatility prediction modes which we apply in our work, the relative error on the fourth moment had the following characteristics: median 0.003%, 99\(^{th}\) percentile 0.105%, maximum 1.659% across 973 observations while we constrained the lattice size $m$ to be no larger than 201.\(^{30}\)

**Appendix C: Trading policy**

We consider calls with moneyness ($K/F$) within the range 0.96-1.08 and puts within the range 0.92-1.04. If we observe $n$ call bid prices violating the call upper bound, each with different strike price, then the option trader writes $1/n$ calls of each type with the underlying futures

\(^{29}\) Note that the presented adjustment of the probabilities in the right tail may not yield an admissible solution, i.e. we may end up with some negative probabilities. If this is the case, we introduce an analogous adjustment in the left tail of the distribution.

\(^{30}\) This lattice size appears unattractive to derive recursive conditional expectations. However, the use of fast Fourier transforms results in a fairly short processing time. See Cerny (2004).
corresponding to the index value of \( y_0 \). The trader transfers the proceeds to the bond account:

\[
x = x_0 + \sum_{i=1}^{n} c_i / n \quad \text{and} \quad y = y_0.
\]

If we observe \( n \) put ask prices violating the put lower bound, each with different strike price, the option trader buys \( 1/n \) puts of each type and finances the purchase out of the bond account: \( x = x_0 - \sum_{i=1}^{n} p_i / n \) and \( y = y_0 \).

However, when there is a violation of the upper put bound and the option trader writes puts, the trader also sells one futures contract for each written put. The intuition for this policy may be gleaned from the observation that the combination of a written put and a short futures amounts to a synthetic short call. In fact, the upper put bound in equation (3) is derived from the upper call bound in equation (2) through the observation that if we can write a put at a sufficiently high price we violate the upper call bound by writing a synthetic call.\(^{31}\)

Finally, when there is a violation of the lower call bound and the option trader buys calls, the trader also sells one futures contract for each purchased call. The intuition is the same as above.

The early exercise policy of a call is based on the function \( N \) in equation (2). However, whenever the option trader is short an option, each period we derive the function \( N \) based on the forward-looking distribution of daily returns, i.e. this function is derived under the empirical distribution of the daily index returns between the option trade and the option maturity. Effectively, we endow the counterparty of the option trader with information on the 2\(^{nd}\), 3\(^{rd}\), and 4\(^{th}\) moments of the forward distribution, while imposing the first moment. The early exercise

\(^{31}\) In implementing the trading policy of either writing puts or buying calls, the option trader buys or sells a futures contract as well and this violates the assumption made in Section I that the option trader does not trade in futures. This, however, is innocuous because, in practice, traders manage their portfolio by trading in the index because of the inconvenience and cost of the frequent rolling over of short-term futures contracts and the illiquidity of long-term futures and forward contracts.
policy of a call or put is simplified by the observation that the decision is a function only of time and the ratio of the strike price to the index level.

Appendix D: The Davidson-Duclos (2000, 2007) tests

The sample counterpart of conditions (4) and (5), applied to the two distributions drawn from their respective populations, is that we must have for every \( z \) in the joint support:

\[
D_{IT}^2(z) - D_{OT}^2(z) > 0, \quad \text{(D.1)}
\]

where

\[
D_j^2(z) = \frac{1}{N} \sum_{i=1}^{N} (z - W_{ji}), \quad \text{(D.2)}
\]

\( N \) is the number of paired outcomes, \( W_{ji} \) is the \( i^{th} \) outcome of the sample \( J \), and \( (x)_+ = \max(x, 0) \).\(^{32}\) Clearly, if (D.1) is violated at any point in the interior of the joint support, the null of non-dominance cannot be rejected. On the other hand, (D.1) becomes, by definition, equality at one or both endpoints of the support. The DD (2006) test deals with this problem by restricting the set of points over which (D.1)-(D.2) are estimated.

DD (2000) provide a test of the null hypothesis \( H_0 : OT \succ IT \) in terms of the maximal and minimal values of the extremal test statistic \( \hat{T}(z) \), defined below. The null is not rejected, if the maximal value of the statistic is positive and statistically significant and the minimal value of the statistic is either positive or negative and statistically not significant. As opposed to DD

\(^{32}\) See DD (2000) for further details.
(2006), this test may provide evidence for stochastic dominance even if we observe a negative statistic \( \hat{T}(z) \).

The variable \( z \) denotes the annualized arithmetic return of a trader, where the subscripts \( IT \) and \( OT \) distinguish between the index trader and the option trader. The statistic \( \hat{T}(z) \) is defined as follows:

\[
\hat{T}(z) = \frac{\hat{D}_{IT}^2(z) - \hat{D}_{OT}^2(z)}{\sqrt{\hat{V}^2(z)}},
\]

where the numerator is given by (D.1)-(D.2) and

\[
\hat{V}^2(z) = \hat{V}_{IT}^2(z) + \hat{V}_{OT}^2(z) - 2\hat{V}_{IT,OT}^2(z)
\]

where

\[
\hat{V}_{IT}^2(z) = \frac{1}{N} \left[ \frac{1}{N} \sum_{i=1}^{N} (z - W_{IT}^i)^2 - \hat{D}_{IT}^2(z) \right], \quad I = IT, OT
\]

and

\[
\hat{V}_{OT,IT}^2(z) = \frac{1}{N} \left[ \frac{1}{N} \sum_{i=1}^{N} (z - W_{IT}^i) (z - W_{OT}^i) - \hat{D}_{IT}^2(z) \hat{D}_{OT}^2(z) \right].
\]

The maximal and minimal values of the statistic are calculated as a maximum and minimum of (D.3) over a set of points of \( z \), as explained below. Stoline and Ury (1979) provide tables for the non-standard distribution of the maximal and minimal value of \( \hat{T}(z) \) at the 1%, 5%, and 10% levels. In principle, the number of points in this joint support over which the test may be performed needs to be restricted since a ‘large’ number of these points violate the independence
assumption between the $\hat{T}(z)$s. Therefore, we compute these statistics for 20 points, equally spaced in the joint support of $W_{IT}$ and $W_{OT}$ (including the endpoints) which corresponds to $k = 20$ in the Stoline and Ury (1979) tables.

By contrast, DD (2006) develop the concept of restricted stochastic dominance in testing the null hypothesis $H_0: OT \not\succ IT$. The test derives the minimal $\hat{T}(z)$-statistic over a suitably restricted interval in the joint support for $IT$ and $OT$. The restriction for the testing interval comes from the observation that a minimal $\hat{T}(z)$-statistic may not be significant by any distributional standards in the tails of the distribution, be it a sample or a population.\footnote{It can be easily shown that the leftmost $T$-statistic is approximately equal to 1, by construction. The numerator of the rightmost $T$-statistic is simply given by (D.1) the difference of the sample means, which implies that testing for SSD at the largest observed outcome corresponds to testing for the significance in the difference in the sample means; this condition is much stronger than necessary for SSD.} Having derived the minimal $\hat{T}(z)$-statistic in a restricted interval, the DD (2006) test applies a bootstrap procedure to the entire data to derive the $p$-value for the test as described below.

A necessary condition for applying DD (2006) is that condition (D.1) holds for our sample. By our trading strategies, condition (D.1) holds over the left side of the return distribution. Its validity, therefore, needs to be tested only over the right side, in which case it corresponds to the positivity of the difference of the means of the two samples. We verify this positivity in all cases and, wherever it is satisfied in the sample, we subject it to further verification by block-bootstrapping 10 years of results from our data. In almost all cases the bootstrap results confirm the sign of the means’ difference.

The test statistic $\hat{T}(z)$ is the same as in DD (2000) and is given by (D.3)-(D.6). This statistic is computed for the values of $z$ that are sample points within the restricted interval, i.e., in this interval we have coupled observations of $W_{IT}$ and $W_{OT}$, transformed to annualized
arithmetic returns. As opposed to the DD (2000) test, there is no restriction on the number of these points and we compute the minimal value of $\hat{T}(z)$ in the restricted interval.\(^{34}\) If the minimal value is negative anywhere in the full support, the null of non-dominance is accepted with p-value of 1.\(^{35}\) Otherwise, we apply the bootstrap approach for the derivation of the \(p\)-values for the null hypothesis, as described in detail in DD (2006) and Davidson (2007). These are simply the number of cases for which the minimal $\hat{T}(z)$ under the bootstrap distribution exceeds the minimum $\hat{T}(z)$ of our sample divided by the number of bootstraps.\(^{36}\) In our tests, we use 999 bootstrap replications in order to derive the \(p\)-values in the tables.

There is a cost in adopting the DD (2006) null, because, as it can be analytically shown, this null cannot be rejected over the entire support of the sample distribution. DD (2006) overcome this problem by restricting the interval over which the null may be rejected to the interior of the support, excluding points at the edges. They then show by simulation that inference on the basis of this restricted interval constitutes the most powerful available inference on the existence of stochastic dominance. In the case of correlated (coupled) samples, the procedure for restricting the interval in the right tail is to start by trimming two pairs of data points: one with the maximal \(W_{IT}\) and the corresponding \(W_{OT}\), and one with the maximal \(W_{OT}\) and the corresponding \(W_{IT}\). We continue in a similar way until the desired degree of trimming is reached. An analogous procedure is implemented in the left tail. Note that the DD (2006) test

\(^{34}\) It may be shown that $\hat{T}(z)$ is monotonic between the sample points; therefore the minimal value of $\hat{T}(z)$ may be found only at a sample point.

\(^{35}\) In this regard, the DD (2006) test is more conservative than the DD (2000) test. The latter test verifies whether an extreme negative $\hat{T}(z)$ is significant for the null $H_0 : IT \succ OT$ while the former test accepts the null of non-dominance.

\(^{36}\) The bootstrap procedure samples all observed coupled values of \(W_{IT}\) and \(W_{OT}\) under artificial probabilities derived for the empirical likelihood maximization under the condition that that the $T(z)$ is set equal to zero at the sample point at which it attains its minimum. See Davidson (2007) for further details.
results for such a procedure are more conservative than those resulting from trimming pairs of observations in the extremes of the tails of the distribution, irrespective of the sample (OT or IT) to which this extreme belongs.

In the on-line appendix of the paper, we verify that trimming on the right does not bias the test towards rejection of the null. In simulated data with characteristics that mirror our sample, we compute the rejection probabilities of the null hypothesis when it is true as well as when it is false. DD (2006) is a weak test without trimming, since it has very low probabilities of rejection of the non-dominance null even when it is false. With 5% trimming, the test is still conservative as far as rejecting the false non-dominance null. Problems with rejection of the null when it is true occur only for deep out-of-the-money options. We control this in our data by restricting the degree of moneyness $K/F$ of the violating options in our tests and show that the null hypothesis $H_0 : OT \not\succ IT$ is again decisively rejected in all cases in spite of the reduced size of the violating sample.
References


Table I
Prediction Error of Monthly Volatility, 1983-2006

<table>
<thead>
<tr>
<th>Prediction mode</th>
<th>Mean</th>
<th>Median</th>
<th>St. dev.</th>
<th>Skew.</th>
<th>Ex. Kurt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td>0.0429</td>
<td>0.0649</td>
<td>0.0680</td>
<td>-1.7300</td>
<td>3.8296</td>
</tr>
<tr>
<td>90-day</td>
<td>0.0095</td>
<td>0.0076</td>
<td>0.0595</td>
<td>0.2687</td>
<td>5.2490</td>
</tr>
<tr>
<td>Adjusted IV</td>
<td>-0.0005</td>
<td>0.0002</td>
<td>0.0496</td>
<td>-0.2625</td>
<td>3.4680</td>
</tr>
<tr>
<td>EGARCH</td>
<td>0.0177</td>
<td>0.0185</td>
<td>0.0531</td>
<td>0.0936</td>
<td>7.8302</td>
</tr>
</tbody>
</table>

The errors are defined as the difference between the monthly volatility and the volatility predicted by a given mode. The unconditional volatility is the sample standard deviation over the period January 1928 to January 1983. The 90-day volatility is the sample standard deviation over the preceding 90 trading days. The adjusted IV is the ATM IV on the preceding day, adjusted by the mean prediction error for all dates preceding the given date, where we drop from the preceding days all 21 pre-crash observations. The EGARCH volatility is the volatility using EGARCH coefficients estimated for S&P 500 daily returns over January 1928 to January 1983 and applied to residuals observed over the 90 days preceding each sample date to form projections of the volatility realized till the option expiration date.

Table II
Percentage of Call Quotes with Violations of the Upper Bound

<table>
<thead>
<tr>
<th>Volatility prediction mode</th>
<th>Moneyness (K/F) Range</th>
<th>0.96-0.99</th>
<th>0.99-1.01</th>
<th>1.01-1.03</th>
<th>1.03-1.08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td></td>
<td>3.3</td>
<td>5.8</td>
<td>8.6</td>
<td>20.6</td>
</tr>
<tr>
<td>90-day</td>
<td></td>
<td>4.8</td>
<td>16.1</td>
<td>32.2</td>
<td>45.4</td>
</tr>
<tr>
<td>IV Adjusted</td>
<td></td>
<td>4.1</td>
<td>15.2</td>
<td>30.5</td>
<td>30.3</td>
</tr>
<tr>
<td>EGARCH</td>
<td></td>
<td>2.3</td>
<td>6.5</td>
<td>14.1</td>
<td>24.3</td>
</tr>
</tbody>
</table>

This table displays the percentages of call bids violating the call upper bound out of all bid quotes observed in each respective moneyness bracket.

Table III
Percentage of Call Quotes with Violations of the Upper Bound out of the Total Number of Quotes

<table>
<thead>
<tr>
<th>Volatility prediction mode</th>
<th>#months with viol. (#months)</th>
<th>Moneyness (K/F) Range</th>
<th>0.96-0.99</th>
<th>0.99-1.01</th>
<th>1.01-1.03</th>
<th>1.03-1.08</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconditional</td>
<td>43 (247)</td>
<td></td>
<td>0.4</td>
<td>1.7</td>
<td>3.5</td>
<td>3.8</td>
<td>9.4</td>
</tr>
<tr>
<td>90-day</td>
<td>100 (247)</td>
<td></td>
<td>0.5</td>
<td>4.8</td>
<td>13.2</td>
<td>8.4</td>
<td>26.9</td>
</tr>
<tr>
<td>IV Adjusted</td>
<td>120 (226)</td>
<td></td>
<td>0.3</td>
<td>4.6</td>
<td>12.9</td>
<td>5.9</td>
<td>23.7</td>
</tr>
<tr>
<td>EGARCH</td>
<td>65 (247)</td>
<td></td>
<td>0.3</td>
<td>1.9</td>
<td>5.8</td>
<td>4.5</td>
<td>12.5</td>
</tr>
</tbody>
</table>

This table displays the percentages of call bids violating the call upper bound in each respective moneyness bracket out of the total number of observed bid quotes.
Table IV
Patterns of Bounds Violations

<table>
<thead>
<tr>
<th>Sort based on:</th>
<th>Volatility Prediction Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unconditional</td>
</tr>
<tr>
<td></td>
<td>Bottom Tercile</td>
</tr>
<tr>
<td><strong>Option Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>mean K/F</td>
<td>0.042</td>
</tr>
<tr>
<td>ATM IV</td>
<td>0.000</td>
</tr>
<tr>
<td>call volume</td>
<td>0.109</td>
</tr>
<tr>
<td>put volume</td>
<td>0.110</td>
</tr>
<tr>
<td>OTM put volume</td>
<td>0.103</td>
</tr>
<tr>
<td>call open interest</td>
<td>0.130</td>
</tr>
<tr>
<td>put open interest</td>
<td>0.139</td>
</tr>
<tr>
<td>OTM put open interest</td>
<td>0.138</td>
</tr>
<tr>
<td>put-call ratio</td>
<td>0.150</td>
</tr>
<tr>
<td>put-call ratio, open interest</td>
<td>0.228</td>
</tr>
<tr>
<td><strong>Index Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>return</td>
<td>0.244</td>
</tr>
<tr>
<td>momentum</td>
<td>0.045</td>
</tr>
<tr>
<td>negative jumps&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.095</td>
</tr>
<tr>
<td>positive jumps&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.109</td>
</tr>
<tr>
<td>yield</td>
<td>0.172</td>
</tr>
<tr>
<td>left skew</td>
<td>0.121</td>
</tr>
<tr>
<td>right skew</td>
<td>0.042</td>
</tr>
<tr>
<td>futures volume</td>
<td>0.090</td>
</tr>
<tr>
<td>futures open interest</td>
<td>0.066</td>
</tr>
<tr>
<td><strong>Interest Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>interest rate</td>
<td>0.150</td>
</tr>
<tr>
<td>yield slope</td>
<td>0.107</td>
</tr>
<tr>
<td>Moody’s Baa-Aaa</td>
<td>0.033</td>
</tr>
</tbody>
</table>

The classified variable is the ratio of the quotes in violation of the call upper bound to the overall number of call bid quotes observed in a given cross-section. The number of cross-sections is 247 for all volatility prediction modes except for Adjusted IV where it is 226. The overall means (st. deviations) of the classified variable for each volatility prediction mode are as follows: Unconditional 0.13 (0.32), 90-day 0.22 (0.35), Adjusted IV 0.24 (0.32), EGARCH 0.13 (0.29). The classifying variables per cross-section are as follows:

Option characteristics: mean K/F is the average moneyness of call bid quotes, ATM IV is the average implied volatility of several bid and ask quotes for calls and puts closely bracketing K/F ratio of 1, call and put volume (open interest) is the natural logarithm of the respective volume (open interest) recorded on the previous day net of the natural logarithm of the average of the respective volume (open interest) over the past 90 calendar days, put-call ratio or similar ratio for open interest was recorded on the previous day.

Index characteristics: return is the S&P 500 index excess return over the previous month, momentum is the ratio of the previous day S&P 500 index level to its average level over the past year, negative (positive) jumps represent the number of the S&P 500 daily returns lower than -4% (greater than 4%) recorded over the past 30 calendar days (it was found 15 (16) occurrences for negative (positive) jumps in total), yield is the ratio of dollar dividends over the past year to the previous day S&P 500 index level, left (right) skew is the implied volatility for the K/F ratio of 0.96 (1.04), futures volume (open interest) is the natural logarithm of the futures volume (open interest) recorded on the previous day net of the natural logarithm of the average of the futures volume (open interest) over the past 90 calendar days.
Interest characteristics: interest rate is the three-month T-bill rate, yield slope is the difference between the yield of 10- and 1-year to maturity CRSP indices for government bonds, Moody’s Baa-Aaa is the difference between the baskets of Baa and Aaa ranked corporate bonds.

Notes:
(a) Instead for terciles, the classified variable was split for dates without and with jumps in the preceding 30 calendar days.
### Table V
Returns of Call Trader and Index Trader

<table>
<thead>
<tr>
<th>Volatility prediction mode</th>
<th>#months with viol. (# months)</th>
<th>$\hat{\mu}<em>{OT} - \hat{\mu}</em>{IT}$ (annualized)</th>
<th>DD (2000) $p$-value</th>
<th>DD (2006) $p$-value</th>
<th>$H_0: OT \succ_2 IT$ 10% trimming in left tail, trimming in right tail as below:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>no trimming</td>
</tr>
<tr>
<td>A: Call Upper Bound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unconditional 43 (247) 0.0031 &lt;0.01 0.244 0.024 0.000 90-day 100 (247) 0.0043 &lt;0.01 0.166 0.007 0.002 Adjusted IV 120 (226) 0.0066* &lt;0.01 0.119 0.029 0.000 EGARCH 65 (247) 0.0062** &lt;0.01 0.079 0.000 0.000</td>
</tr>
<tr>
<td>B: Call Upper Bound with 20% of HML and 20% of SMB in Stock Account</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unconditional 43 (247) 0.0031 &lt;0.05 0.244 0.126 0.038 90-day 100 (247) 0.0043 &lt;0.01 0.149 0.056 0.012 Adjusted IV 120 (226) 0.0066* &lt;0.01 0.133 0.052 0 EGARCH 65 (247) 0.0062** &lt;0.01 0.072 0.006 0</td>
</tr>
<tr>
<td>C: Put Upper Bound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unconditional 23 (247) 0.0009 &gt;0.1 0.399 0.203 0.154 90-day 16 (247) -0.0008 &gt;0.1 1 1 1 Adjusted IV 4 (226) n/a n/a n/a n/a n/a EGARCH 9 (247) n/a n/a n/a n/a n/a</td>
</tr>
</tbody>
</table>

Equally weighted average of all violating options equivalent to one option per share was traded at each date. The symbols * and ** denote a difference in sample means of the OT and IT traders significant at the 5% and 1% levels in a one-sided bootstrap test with 9,999 trials. Maximal $t$-statistics for Davidson-Duclos (DD, 2000) test are compared to critical values of Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for three nominal levels of 1, 5, and 10% with $k = 20$ and $\nu = \infty$. The $p$-values for $H_0: OT \succ_2 IT$, which are greater than 10%, the highest nominal level available in Stoline and Ury (1979) tables are not reported here. P-values for the Davidson-Duclos (2007) test are based on 999 bootstrap trials. The $p$-values for $H_0: IT \succ_2 OT$ are equal to one and are not reported here. HML and SMB denote Fama-French factors, i.e. HML denotes a portfolio long in two value portfolios and short in two growth portfolios and SMB denotes a portfolio long in three small-cap portfolios and short in three large-cap portfolios. See Fama and French (1993) for details.
### Table VI

Returns of Options Trader and Index Trader—without Futures Basis Risk

<table>
<thead>
<tr>
<th>Volatility prediction mode</th>
<th>#months with viol. (# months)</th>
<th>( \hat{\mu}<em>{OT} - \hat{\mu}</em>{IT} ) (annualized)</th>
<th>DD (2000) ( p )-value ( H_0 : IT \gtrsim \ot )</th>
<th>DD (2006) ( p )-value ( H_0 : OT \gtrsim IT )</th>
</tr>
</thead>
</table>

#### A: Call Upper Bound

<table>
<thead>
<tr>
<th></th>
<th>Unconditional</th>
<th>90-day</th>
<th>Adjusted IV</th>
<th>EGARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>67 (247)</td>
<td>156 (247)</td>
<td>195 (226)</td>
<td>112 (247)</td>
</tr>
<tr>
<td>( \hat{\mu}<em>{OT} - \hat{\mu}</em>{IT} ) (annualized)</td>
<td>0.0012</td>
<td>0.0083**</td>
<td>0.0032</td>
<td>0.0037</td>
</tr>
<tr>
<td>( H_0 : IT \gtrsim \ot )</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DD (2000) ( p )-value</td>
<td>0.412</td>
<td>0.083</td>
<td>0.337</td>
<td>0.261</td>
</tr>
<tr>
<td>DD (2006) ( p )-value</td>
<td>0.128</td>
<td>0.011</td>
<td>0.255</td>
<td>0.074</td>
</tr>
</tbody>
</table>

#### B: Put Upper Bound

<table>
<thead>
<tr>
<th></th>
<th>Unconditional</th>
<th>90-day</th>
<th>Adjusted IV</th>
<th>EGARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>36 (247)</td>
<td>52 (247)</td>
<td>64 (226)</td>
<td>38 (247)</td>
</tr>
<tr>
<td>( \hat{\mu}<em>{OT} - \hat{\mu}</em>{IT} ) (annualized)</td>
<td>-0.0015</td>
<td>-0.0003</td>
<td>-0.0012</td>
<td>0.0014</td>
</tr>
<tr>
<td>( H_0 : IT \gtrsim \ot )</td>
<td>&lt;0.1</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>DD (2000) ( p )-value</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DD (2006) ( p )-value</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The table differs from Table V only in that the basis risk is set at zero, \( \bar{\varepsilon} = 0 \), instead of bounding the risk by \( \bar{\varepsilon} = 0.5\% \). Equally weighted average of all violating options equivalent to one option per share was traded at each date. The symbol ** denotes a difference in sample means of the \( OT \) and \( IT \) traders significant at the 5% level in a one-sided bootstrap test with 9,999 trials. Maximal \( t \)-statistics for Davidson-Duclos (DD, 2000) test are compared to critical values of Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for three nominal levels of 1, 5, and 10% with \( k = 20 \) and \( \nu = \infty \). The \( p \)-values for \( H_0 : OT \gtrsim IT \), which are greater than 10%, the highest nominal level available in Stoline and Ury (1979) tables are not reported here. \( p \)-values for the Davidson-Duclos (2007) test are based on 999 bootstrap trials. The \( p \)-values for \( H_0 : IT \gtrsim OT \) are equal to one and are not reported here.
Table VII
Returns of Naïve Options Trader and Index Trader

<table>
<thead>
<tr>
<th>Trade type or Volatility Est. Mode</th>
<th>#months with viol. (# months)</th>
<th>$\hat{\mu}<em>{OT} - \hat{\mu}</em>{IT}$ (annualized)</th>
<th>DD (2000) $p$-value $H_0 : IT \neq_2 OT$</th>
<th>DD (2006) $p$-value $H_0 : OT \neq_2 IT$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: All Quantiles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short call</td>
<td>247 (247)</td>
<td>0.0060</td>
<td>&lt;0.01</td>
<td>0.207</td>
</tr>
<tr>
<td>Short put</td>
<td>247 (247)</td>
<td>0.0078</td>
<td>&lt;0.01</td>
<td>0.280</td>
</tr>
<tr>
<td>Long call</td>
<td>247 (247)</td>
<td>-0.0403***</td>
<td>&lt;0.01</td>
<td>1</td>
</tr>
<tr>
<td>Long put</td>
<td>247 (247)</td>
<td>-0.0292***</td>
<td>&lt;0.01</td>
<td>1</td>
</tr>
<tr>
<td>B: 10th or 90th Critical Quantile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short call</td>
<td>58 (243)</td>
<td>0.0041</td>
<td>&lt;0.01</td>
<td>0.195</td>
</tr>
<tr>
<td>Short put</td>
<td>67 (243)</td>
<td>0.0034</td>
<td>&lt;0.01</td>
<td>0.331</td>
</tr>
<tr>
<td>Long call</td>
<td>73 (243)</td>
<td>-0.0149***</td>
<td>&gt;0.1</td>
<td>1</td>
</tr>
<tr>
<td>Long put</td>
<td>95 (243)</td>
<td>-0.0058</td>
<td>&gt;0.1</td>
<td>1</td>
</tr>
<tr>
<td>C: 2.5th or 97.5th Critical Quantile</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short call</td>
<td>32 (243)</td>
<td>0.0057</td>
<td>&lt;0.01</td>
<td>0.073</td>
</tr>
<tr>
<td>Short put</td>
<td>36 (243)</td>
<td>0.0022</td>
<td>&gt;0.1</td>
<td>0.359</td>
</tr>
<tr>
<td>Long call</td>
<td>27 (243)</td>
<td>-0.0038</td>
<td>&gt;0.1</td>
<td>1</td>
</tr>
<tr>
<td>Long put</td>
<td>45 (243)</td>
<td>+0.0000</td>
<td>&gt;0.1</td>
<td>0.482</td>
</tr>
</tbody>
</table>

Equally weighted average of all violating options equivalent to one option per share was traded at each date. The symbols *, **, and *** denote a difference in sample means of the OT and IT traders significant at the 10, 5 and 1% level in a one-sided bootstrap test with 9,999 trials. Maximal $t$-statistics for Davidson-Duclos (DD, 2000) test are compared to critical values of Studentized Maximum Modulus Distribution tabulated in Stoline and Ury (1979) for three nominal levels of 1, 5, and 10% with $k = 20$ and $\nu = \infty$. The $p$-values for $H_0 : OT \neq_2 IT$, which are greater than 10%, the highest nominal level available in Stoline and Ury (1979) tables are not reported here. $p$-values for the Davidson-Duclos (2007) test are based on 999 bootstrap trials. The $p$-values for $H_0 : IT \neq_2 OT$ are equal to one and are not reported here.
Figure 1: Illustration of Upper and Lower Bounds on Call and Put Options

Bounds are derived for $\sigma = 0.20$ imposed on a 90-day past distribution of daily S&P 500 returns for May 22, 1996. 95% upper and lower confidence intervals represented by dotted lines are derived by bootstrapping the 90-day distribution. The results exemplify the dependence of the bounds on the third and fourth moments of the distribution because the width in of the confidence intervals is determined solely by varying the skewness and kurtosis, i.e. the bootstrap changes only these quantities.
The crosses display the violations of the call upper bound for the period February 1983-July 2006. For the adjusted IV distribution, the first 21 dates are not in the sample. To facilitate presentation, the S&P Index was transformed to a logarithmic scale. The inception date of the VIX index was on February 4th, 1986. The value for VIX just prior to the October 1987 crash was 170% and is trimmed to facilitate presentation.